Linkage of Early-Onset Familial Breast Cancer to Chromosome 17q21
Author(s): Jeff M. Hall, Ming K. Lee, Beth Newman, Jan E. Morrow, Lee A. Anderson, Bing Huey, Mary-Claire King
Reviewed work(s):
Source: *Science*, New Series, Vol. 250, No. 4988 (Dec. 21, 1990), pp. 1684-1689
Published by: American Association for the Advancement of Science
Stable URL: http://www.jstor.org/stable/2878541
Accessed: 03/12/2011 17:28

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at http://www.jstor.org/page/info/about/policies/terms.jsp

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

American Association for the Advancement of Science is collaborating with JSTOR to digitize, preserve and extend access to *Science*.

http://www.jstor.org

Linkage of Early-Onset Familial Breast Cancer to Chromosome 17q21

JEFF M. HALL, MING K. LEE, BETH NEWMAN, JAN E. MORROW, LEE A. ANDERSON, BING HUEY, MARY-CLAIRE KING

Human breast cancer is usually caused by genetic alterations of somatic cells of the breast, but occasionally, susceptibility to the disease is inherited. Mapping the genes responsible for inherited breast cancer may also allow the identification of early lesions that are critical for the development of breast cancer in the general population. Chromosome 17q21 appears to be the locale of a gene for inherited susceptibility to breast cancer in families with early-onset disease. Genetic analysis yields a lod score (logarithm of the likelihood ratio for linkage) of 5.98 for linkage of breast cancer susceptibility to *D17S74* in early-onset families and negative lod scores in families with late-onset disease. Likelihood ratios in favor of linkage heterogeneity among families ranged between 2000:1 and greater than 10^6:1 on the basis of multipoint analysis of four loci in the region.

HUMAN DISEASE GENES CAN BE LOCATED BY LINKAGE analysis of families in which the incidence of the disease is high. Linkage analysis can reveal the chromosomal location of the genes of interest by identifying polymorphic genetic markers of known location that are coinherited with the disease in families (*1*). Among the common cancers, breast cancer is particularly suited for this approach, because family history of the disease is a significant risk factor in all populations; epidemiological evidence consistently indicates that a woman's risk of breast cancer is increased by the occurrence of the disease in her mother or sisters. The younger the ages at diagnosis of her relatives, the greater the increase in a woman's risk (*2*).

The transformation of breast ductal epithelial cells to malignant growth results from alterations in their DNA that may be either inherited or somatic (*3*). Mapping genes for familial breast cancer is important because alterations at the same loci may also be responsible for sporadic disease. Individuals with inherited susceptibility to breast cancer are completely asymptomatic for decades before the onset of disease; the effects of critical inherited alterations are thus latent for an extended period. Among women with no inherited susceptibility to the disease, these same alterations may be the initial lesions of breast tumorigenesis, with disease expression being similarly dependent on subsequent genetic alterations or tumor-promoting steps.

Mapping genes for human breast cancer has been complicated by unavoidable epidemiologic realities. The disease is common, but only a small proportion of cases in the general population are attributable to inherited susceptibility. Thus, families may have multiple cases of breast cancer without inherited susceptibility, and "sporadic" cases may occur even in families with inherited disease. In addition, the disease is not completely penetrant among susceptible persons, with expression depending on gender, age, and nongenetic risk factors. Finally, both epidemiological and molecular evidence suggests heterogeneity. We have tested simultaneously for genetic linkage and heterogeneity of breast cancer in families, and our results suggest both the presence of a gene for early-onset breast cancer on chromosome 17q21 and linkage heterogeneity of the disease.

Families and inheritance of susceptibility. Our genetic analysis is based on 23 extended families with 146 cases of breast cancer (Figs. 1 and 2). All persons in our analysis are Caucasian and from a variety of original ancestries. The 329 participating relatives now live in, and were therefore sampled from, 40 states of the United States, Puerto Rico, Canada, the United Kingdom, and Colombia. These families share the epidemiological features that are characteristic of familial, versus sporadic, breast cancer (*2*): younger age at diagnosis, frequent bilateral disease, and more frequent occurrence of disease among men.

Our statistical model for the inheritance of susceptibility to breast cancer was derived from our previous complex segregation analysis of a population-based series of 1500 families with breast cancer (*4*). Inherited susceptibility to breast cancer in that series could be fully explained by a rare autosomal dominant allele with a major effect on risk: risk of breast cancer in genetically susceptible women was estimated to be 0.37 by age 40, 0.66 by age 55, and 0.82 over the entire lifetime. In contrast, risk of breast cancer in women without genetic susceptibility was estimated to be 0.004 by age 40, 0.028 by age 55, and 0.081 over the entire lifetime. Females less than 15 years of age and all males had a negligible risk (less than 0.001). The estimated proportion of breast cancer cases in the sample that were attributable to inherited susceptibility was only 4 percent, the great majority of cases resulting purely from somatic events. Among younger patients, however, the proportion of inherited cases is likely to be considerably higher. Disease allele frequencies (q) between 0.004 and 0.02 yield virtually identical results; those for q equals 0.01 are described.

Definition of the breast cancer phenotype. For any complex disease, it is essential to adequately define the phenotype, the inheritance of which will be traced in families. Real linkages can be missed and spurious linkages suggested either by defining the phenotype too broadly (so that persons without inherited susceptibility to disease are mistakenly categorized as affected) or simply by making errors in diagnosis. To minimize errors in diagnosis, we

The authors are at the School of Public Health, University of California, Berkeley, CA 94720.

reviewed existing pathology records of all family members on whom breast surgery had been performed. For deceased persons reported by their relatives to have had breast cancer, but for whom no pathology records were available, we obtained hospital records or death certificates. For living subjects who had not undergone breast surgery, we relied on self-report of no breast cancer; for deceased persons with no history of breast surgery, we relied on death certificates and reports of relatives. The affected phenotype was defined as all histologic types of invasive breast cancer. No other cancer sites were included (*5*).

Statistical methods. Four approaches to evaluating linkage were applied:

1) Lod scores (logarithms of the likelihood ratios for linkage) for linkage of individual markers to disease were estimated with the LIPED program; multipoint mapping was performed with the LINKAGE program; homogeneity of recombination fractions was evaluated with the B test for two-point linkage data and direct comparison of likelihoods for multipoint data (*1, 6*). Homogeneity of linkage was tested for all families; the sample was not a priori subdivided. Linkage analyses were based on an autosomal dominant model with the age- and sex-specific risks described above for hypothetically susceptible and nonsusceptible individuals. (LIPED was modified to incorporate four liability classes for each genotype.) The risk group (liability class) for each woman was defined by her age at diagnosis of breast cancer, at death if deceased without breast cancer, or at most recent interview if living and unaffected. All men were assigned to the lowest risk group.

2) In determining a plausible model for inheritance of susceptibility to a complex disease, it is always possible that the underlying genetic model may be correct, but that penetrance may not be accurately estimated. Therefore, we also tested for linkage by including disease information only for the affected relatives (that is, the breast cancer cases) in each family; all unaffected subjects were assigned to the lowest risk class, so that only their marker information was incorporated. Autosomal dominance was still assumed, and lifetime risk of sporadic disease was not altered.

3) In order to evaluate linkage without imposing any specific genetic model, we counted alleles shared by descent for affected pairs of female relatives. This analysis was possible because one marker was highly polymorphic. Pairs of affected relatives were sisters (first-degree relatives), aunt-niece and grandmother-granddaughter (second-degree relatives), and first cousins (third-degree relatives). Relative pairs were stratified by the average age of breast cancer diagnosis for the pair.

4) We applied the affected-pedigree-member method to evaluate sharing of alleles by state among the breast cancer cases in each family (*7*). Only cases whose marker genotypes could be determined with certainty were included in the analysis of affected pedigree members.

Typing of DNA polymorphisms. For each of the 329 informative relatives, we obtained 35 milliliters of fresh blood and prepared immortalized lymphocytes by Epstein-Barr viral transformation (*8*). Genomic DNA was prepared as described (*9*). Probes were labeled by random primer extension (*10*) and hybridized to DNA according to standard procedures (*11*). Parentage was confirmed by consistent inheritance of 183 polymorphic markers. Markers at chromosome 17q21 include *D17S74*, a VNTR (variable number of tandem repeats) defined by the probe CMM86 and HinfI; *D17S40*, defined by the probe LEW 101 and Msp I; *D17S41*, defined by the probe LEW 102 and Pst I; and *D17S78*, defined by the probe p131 and

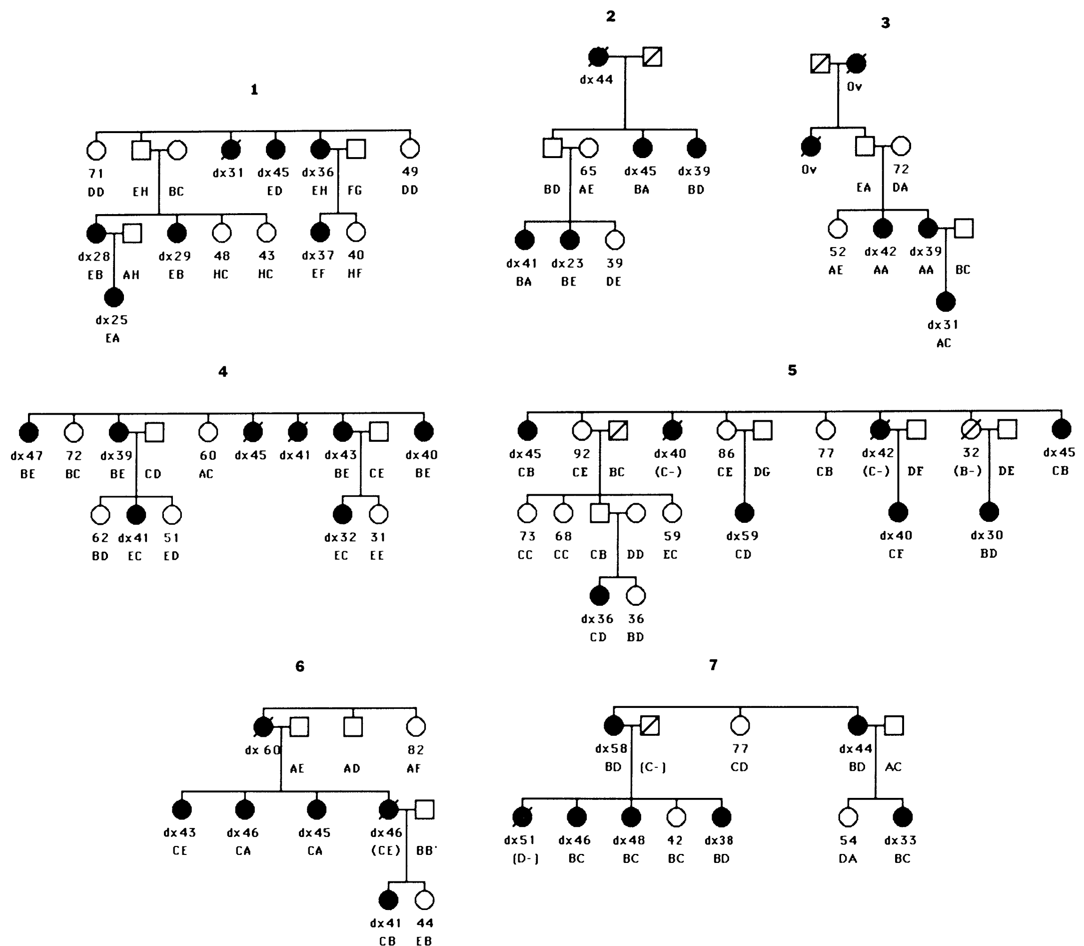

Fig. 1. Breast cancer families 1 to 7. Solid circles, females with breast cancer; open circles, females without breast cancer; open squares, males without breast cancer. The age given for each woman is age at (first) breast cancer diagnosis (dx) if affected, age at death if deceased (deceased individuals are represented by diagonal lines through symbols), or age at most recent interview if alive without breast cancer. Alleles of *D17S74* are shown for all families and are lettered sequentially within each family from largest to smallest fragment size. Alleles in parentheses are based on reconstructed genotypes.

Msp I (12). On the basis of analysis of the CEPH families, the order and approximate recombination distances of the markers are: 17cen—*D17S78*–(0.10)–*D17S41*–(0.06)–*D17S74*–(0.12)–*D17S40* —17qter (13).

D17S74, a highly polymorphic VNTR with heterozygosity greater than 0.90, is extremely useful for linkage analysis but presents technical challenges that are common to VNTRs. Linkage analysis is critically sensitive both to errors in assigning genotypes and to marker allele frequencies. The lengths of the HinfI restriction fragments that define *D17S74* alleles are continuously distributed between 1 and 5 kilobase pairs (kb). For our analyses, *D17S74* alleles were identified by analyzing DNA samples from all members of a family on the same Southern blot, placing relatives with fragments of similar size adjacent to one another (14). "Population"

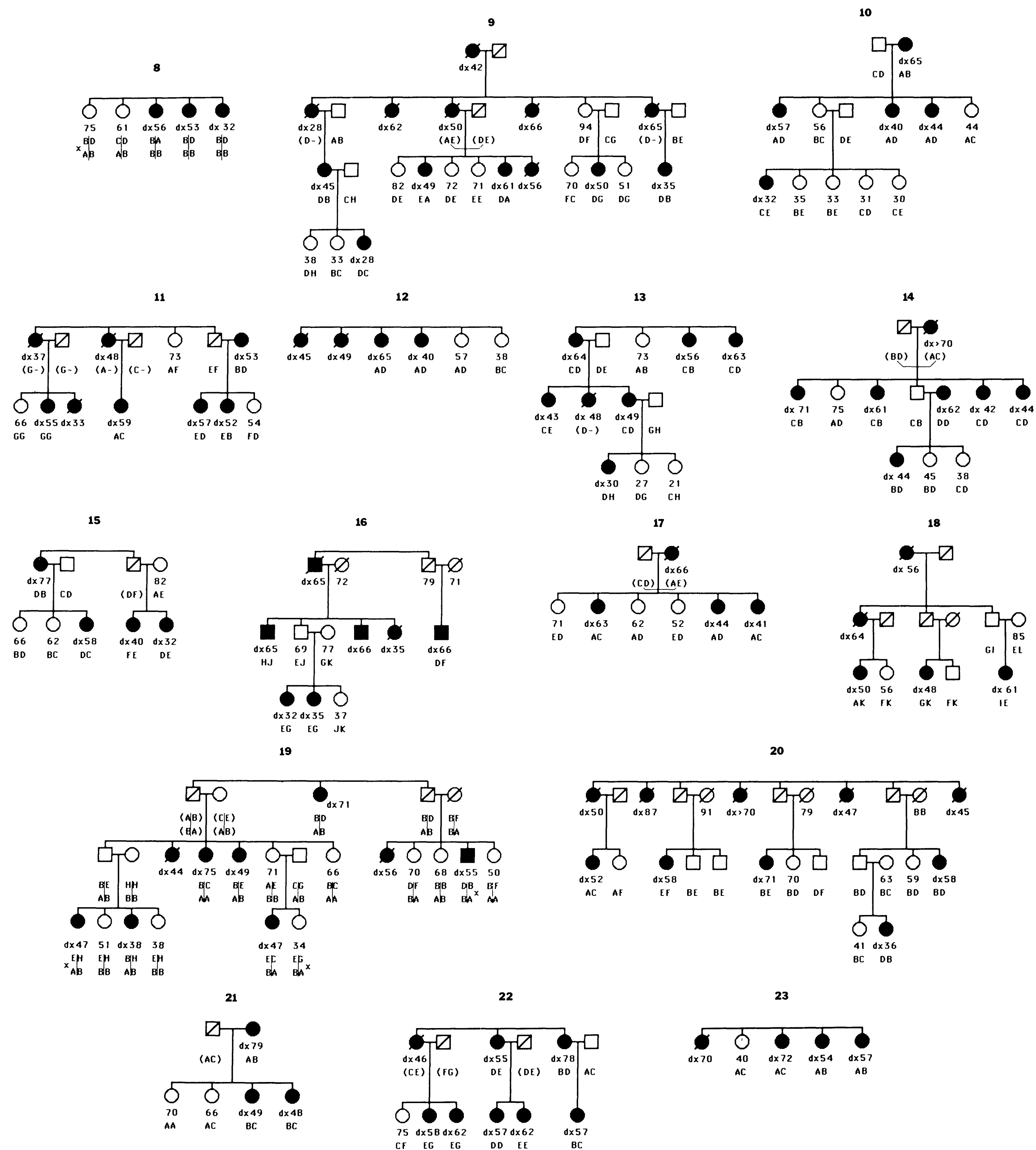

Fig. 2. Breast cancer families 8 to 23. Notation as for Fig. 1, with solid squares in families 16 and 19 representing males with breast cancer. Genotypes for *D17S74* are shown for all families. *D17S78* genotypes for family 8 and *D17S40* genotypes for family 19 are also shown.

frequencies of the *D17S74* alleles in this sample were estimated by selecting subjects from different families whose *D17S74* fragments appeared to be of similar size on the basis of their "family blots," and then analyzing the DNA from these unrelated persons in neighboring lanes on the same blots. Some samples were included several times in order to identify distortions in the gels. These "population blots" were analyzed without reference to sample numbers, in order to determine which alleles could be consistently distinguished. *D17S74* had more than 30 distinguishable fragment lengths—and hence more than 30 different alleles—in our sample, nine of which occurred more than once among unrelated individuals, at frequencies ranging from 0.07 to 0.13. The other *D17S74* alleles were only represented once in our sample, but because extremely rare marker allele frequencies can have a major influence on estimates of lod scores and the *T* statistic (*7*), apparently unique alleles were each assigned the frequency 0.03.

Results of linkage and heterogeneity analysis in the breast cancer families. For the 23 families as a group, homogeneity of linkage of breast cancer to *D17S74* could be rejected at *P* equals 0.01. Multipoint analysis of linkage in the interval *D17S78-D17S41-D17S74-D17S40* yielded likelihood ratios in favor of heterogeneity of linkage among the 23 families between 2000:1 and 1.4×10^6 to 1. After adjusting for heterogeneity among all families, the maximum two-point lod score is +3.28 at recombination distance of 0.014 from *D17S74*, with disease linked to this locus in 40 percent of the families (Figs. 1 and 2).

Heterogeneity of linkage in these families appears explicable by age of disease onset. Breast cancer is linked to markers in this chromosomal region specifically in families with early-onset breast cancer. Among the seven families with a mean age of breast cancer diagnosis less than or equal to 45, the two-point lod score for linkage of *D17S74* and breast cancer is +5.98 at a distance of 0.001 recombination units, with a 95 percent confidence interval of 0.001 to 0.09 (Table 1). In contrast, total lod scores for the families with late-onset disease are negative. It is characteristic of linkage in the presence of heterogeneity that a modest lod score (in this case +2.35) for all families, ignoring heterogeneity, at a fairly large recombination distance (0.20 recombination units) masks two curves, one with a more positive lod score in the linked families (+5.98) at a smaller recombination fraction (0.001 to 0.09) and the other negative (*15*).

Two-point lod scores for all four markers in this chromosomal region suggest that a gene for susceptibility in the early-onset families is likely to be within approximately 10 percent recombination of *D17S74* (Table 2). Multipoint analysis of the four-marker interval yields a maximum lod score of +5.41 near *D17S74* for the earliest-onset families (Table 2). Again, total lod scores for families with older ages at diagnosis are negative throughout the interval.

Linkage of breast cancer to *D17S74* was also evaluated on the basis of only the individuals with breast cancer in each family. For this analysis, all women without breast cancer and all men were assigned to the lowest risk group. For the 23 families as a group, the *P* value for homogeneity of linkage is 0.06. For the families with average age at diagnosis less than or equal to 45, the maximum lod score is +4.69 at close linkage, with a 95 percent confidence interval for the recombination fraction of 0.001 to 0.10. Lod scores at close linkage to *D17S74* are −2.19 for families 8 to 15 and −5.22 for families 16 to 23.

Analysis of alleles shared by descent among related pairs of women with breast cancer also suggested linkage of early-onset breast cancer to *D17S74* (Table 3). In families 1 to 7, all three classes of relatives shared more alleles by descent than expected by chance. Even in families 8 to 23, there was evidence for increased identity by descent among sisters with early-onset disease, but little or no evidence for increased identity by descent for second- and third-degree relatives. Finally, the affected-pedigree-member analysis of alleles shared by state among all individuals with breast cancer in each family also suggested linkage of early-onset breast cancer to markers in this chromosomal region (*16*).

Table 1. Lod scores for linkage of breast cancer to D17S74, chromosome 17q21. For each family, *M* is the mean age of diagnosis of breast cancer.

Family	*M*	Recombination fraction					ΣZ at 0.001
		0.001	0.10	0.20	0.30	0.40	
1	32.7	+2.36	+1.89	+1.38	+0.82	+0.28	+2.36
2	37.2	+0.50	+0.35	+0.21	+0.09	+0.02	+2.86
3	37.3	+0.40	+0.29	+0.19	+0.09	+0.03	+3.26
4	39.8	+1.14	+0.91	+0.64	+0.35	+0.11	+4.40
5	42.6	−0.50	−0.25	−0.08	+0.00	+0.03	+3.90
6	44.2	+1.38	+1.06	+0.73	+0.41	+0.14	+5.28
7	45.4	+0.70	+0.58	+0.40	+0.21	+0.05	+5.98
8	47.0	+0.00	+0.02	+0.02	+0.01	+0.00	+5.98
9	47.4	−0.31	+0.03	+0.06	+0.04	+0.01	+5.67
10	47.6	−0.04	−0.06	−0.08	−0.08	−0.05	+5.63
11	49.3	−1.51	−0.41	−0.13	−0.03	−0.00	+4.12
12	50.2	−0.06	−0.03	−0.02	−0.01	−0.00	+4.06
13	50.4	−0.41	−0.09	−0.02	−0.03	−0.04	+3.65
14	51.4	−0.65	−0.18	+0.01	+0.06	+0.04	+3.00
15	51.8	−0.35	−0.08	−0.02	−0.01	−0.00	+2.65
16	52.0	−2.71	−0.56	−0.20	−0.07	−0.02	−0.06
17	53.5	−0.13	+0.04	+0.07	+0.05	+0.01	−0.19
18	53.6	−0.75	−0.38	−0.18	−0.07	−0.02	−0.94
19	55.8	−2.56	−0.93	−0.45	−0.20	−0.05	−3.50
20	56.4	−1.71	−1.01	−0.56	−0.28	−0.11	−5.21
21	58.7	+0.65	+0.50	+0.34	+0.18	+0.05	−4.56
22	59.4	−0.85	−0.13	+0.04	+0.05	+0.02	−5.41
23	63.3	−0.07	−0.02	+0.00	+0.00	+0.00	−5.48

These analyses defined families with early-onset disease as those with an average age at breast cancer diagnosis of less than or equal to 45. However, "early-onset" could be defined in a variety of ways. The cumulative lod scores for linkage of breast cancer to *D17S74*, as families with increasing age at disease onset are included in the analysis, are indicated in Fig. 3 and the right-hand column of Table 1. Cumulative lod scores are above 5.0 for families with an average age at diagnosis of less than 48, remain positive for families with an average age at diagnosis of less than 52, and then drop sharply. As specific alternative ways of defining early-onset breast cancer, we defined families with early-onset disease to be those in which (i) most breast cancers in each family were diagnosed by age 50 (families 1 to 7, 9, 10, 12, 13, 21), or (ii) most breast cancers in each family were diagnosed by age 45 (families 1 to 6, 10), or (iii) the average age at breast cancer diagnosis was less than 48, the mean for all the cases in the sample (families 1 to 10). The critical results for all definitions were the same: maximum lod scores for linkage of breast cancer to this chromosomal region are between 5.2 and 5.7 in the early-onset families and negative in the late-onset families, and the disease gene appears to be within about 0.10 recombination units of *D17S74*.

Other risk factors for breast cancer in the families. To determine whether the linked gene is expressed in the presence of any specific background of nongenetic risk factors for breast cancer, we compared breast cancer risk factors for women in families with apparent linkage of breast cancer to chromosome 17q versus women in families with evidence against this linkage. Ages at first pregnancy, number of children, prevalence of fertility problems, exposure to x-rays, use of oral contraceptives, and ages at menopause were similarly distributed in "linked" and "unlinked" families, after adjusting for age and birth cohort of the women. Similarly, preva-

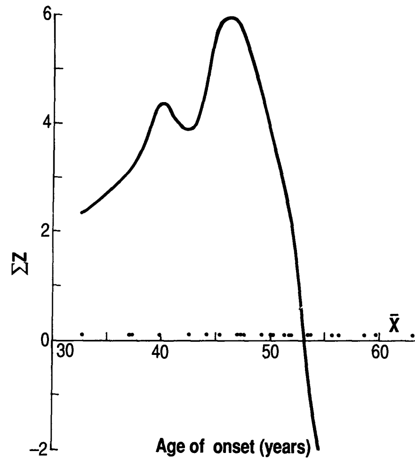

Fig. 3. Linkage of *D17S74* to breast cancer in families, based on the autosomal dominant model described in the test. Cumulative lod scores (ΣZ) are shown for all families for which the mean age of breast cancer diagnosis (M) is less than or equal to the age represented on the x-axis. Total lod scores are greater than or equal to 5.0 for families with $M \leq 48$ and greater than or equal to 3.0 for families with $M \leq 50$. Each dot above the x-axis represents a family with that particular mean onset age.

lences of specific cancers at other sites did not differ among women in "linked" and "unlinked" families, although male breast cancers occurred only in "unlinked" families. The only observed difference between women in "linked" versus "unlinked" families was age at breast cancer diagnosis.

Linkage analysis in families and loss of heterozygosity in tumors. Comparisons of breast tumor tissue with normal tissue from the same individual have suggested that chromosomes 1p, 3p, 11p, 13q, 16q, or 17p (or some combination) may harbor genes that are important for breast tumor progression (*17*). An earlier linkage analysis from our group, based on fewer and less-informative families, suggested (with a modest lod score) that a gene for familial premenopausal breast cancer or ovarian cancer is present on chromosome 16q (*18*). We have excluded the other regions suggested by loss of heterozygosity in tumors as locales of genes for inherited susceptibility to breast cancer (*9, 19*).

A gene or genes on chromosome 5q appear to be responsible for inherited forms of colon cancer and quite possibly for the initial somatic lesions of other colon cancers, with genes on 12p (*K-ras*), 17p (p53), 18q (DCC), and possibly elsewhere being responsible for subsequent invasion and metastasis (*20*). This pattern of multiple sequential events, determined by alterations on more than one chromosome, may also apply to breast cancer.

Negative lod scores in the families with late-onset breast cancers may reflect any or all of the following: the existence of a different locus or loci responsible for inherited susceptibility in these families; the chance occurrence of some families with multiple cases; or a higher prevalence of sporadic cases in families with late-onset inherited disease. Loci responsible for disease in families with late-onset disease may be identifiable by continued simultaneous analysis of linkage and heterogeneity.

Table 3. Identity by descent of two, one, or zero alleles of *D17S74* among pairs of related individuals with breast cancer. Mean onset is the average age of breast cancer diagnosis for the related pair. Family numbers refer to Figs. 1 and 2. Second-degree (2°) relatives are aunt-niece or grandmother-granddaughter pairs; third-degree (3°) relatives are first cousins.

Mean onset of related pair (years)	Sisters			2°		3°	
	Two	One	Zero	One	Zero	One	Zero
			All families				
≤45	14	4	0	19	5	8	4
46–55	5	8	0	9	2	4	4
≥56	3	6	2	7	7	3	10
			Families 1 to 7				
≤45	9	2	0	17	0	6	2
46–55	0	0	0	2	0	1	0
			Families 8 to 23				
≤45	5	2	0	2	5	2	2
46–55	5	8	0	7	2	3	4
≥56	3	6	2	7	7	3	10

Candidate genes on chromosome 17q. The ultimate goal of gene mapping of human traits is to move from a known chromosomal location to identification of the crucial gene and characterization of its critical alterations. This region of chromosome 17q includes several plausible candidate genes (*21*). A gene for a truncated form of the human epidermal growth factor receptor [*her2*; MIM 164870 (Mendelian Inheritance in Man)] is identical to *erbb2* or *neu* (MIM 190150). The gene *her2* acts as an oncogene in NIH 3T3 cells and is amplified in many primary breast tumors; *her2* amplification is associated with poor prognosis at least for node-positive tumors (*22*). Other candidate genes in this region include that for estradiol-17β dehydrogenase (*edhb17*; MIM 264300), which is the enzyme that catalyzes the conversion of estrone to estradiol; the homeobox 2 gene (*hox2*; MIM 142960), which is critical for murine embryological development; *nm23*, whose expression is associated with lymph node metastasis in primary breast carcinomas; the gene for retinoic acid receptor α (*rara*; MIM 180240), a protein that binds the possible anticarcinogen retinoic acid; and *wnt3* (MIM 165330), one of the integration sites activated by the mouse mammary tumor virus and which is homologous to

Table 2. Linkage analysis of breast cancer to four markers on chromosome 17q by mean age (M) of breast cancer diagnosis.

Marker	Two-point lod scores at five recombination fractions: Families 1 to 7 ($M \leq 45$)					Familes 8 to 15 ($46 \leq M \leq 51$)					Families 16 to 23 ($M \leq 52$)				
	0.001	0.10	0.20	0.30	0.40	0.001	0.10	0.20	0.30	0.40	0.001	0.10	0.20	0.30	0.40
D17S78	−0.84	−0.65	−0.16	−0.04	+0.00	+0.36	+0.95	+0.81	+0.48	+0.14	−4.18	−1.25	−0.44	−0.12	−0.04
D17S41	−1.54	−1.12	−0.71	−0.36	−0.14	+0.10	+0.51	+0.43	+0.26	+0.08	−2.73	−1.03	−0.59	−0.35	−0.16
D17S74	+5.98	+4.83	+3.47	+1.97	+0.65	−3.33	−0.80	−0.18	−0.05	−0.04	−8.13	−2.49	−0.94	−0.34	−0.12
D17S40	+1.36	+1.01	+0.63	+0.30	+0.07	−0.49	−0.12	+0.01	+0.06	+0.05	−2.42	−0.71	−0.25	−0.05	−0.00

Families	Lod scores based on multipoint analysis of the interval (on recombination units): *S78* .000	0.020	0.040	0.060	0.080	*S41* 0.100	0.120	0.140	*S74* 0.160	0.184	0.208	0.232	0.256	*S40* 0.280
1 to 7	+2.83	+3.09	+3.30	+3.47	+3.57	+3.41	+4.46	+4.60	+5.24	+5.41	+5.24	+4.96	+4.48	
8 to 15	−0.30	−0.07	+0.01	+0.03	−0.05	−0.20	−1.58	−2.71	−9.14	−5.61	−4.24	−3.36	−2.78	
16 to 23	−6.70	−5.80	−5.51	−5.52	−5.89	−6.98	−6.60	−7.94	−15.21	−8.94	−6.79	−5.58	−5.03	

the *Drosophila* winglessness locus (*23*). If alterations in any of these genes are responsible for inherited breast cancer, polymorphisms at the critical locus may be in linkage disequilibrium with the disease in the early-onset families.

REFERENCES AND NOTES

1. J. Ott, *Analysis of Human Genetic Linkage* (Johns Hopkins Univ. Press, Baltimore, 1985).
2. R. Ottman, M. C. Pike, M.-C. King, J. T. Casagrande, B. E. Henderson, *Am. J. Epidemiol.* **123**, 15 (1986); N. L. Petrakis, V. L. Ernster, M.-C. King, in *Cancer Epidemiology and Prevention*, D. Schottenfeld and J. F. Fraumeni, Jr., Eds. (Saunders, Philadelphia, 1982), pp. 855–870; A. G. Schwartz, M.-C. King, S. H. Belle, W. A. Satariano, G. M. Swanson, *J. Natl. Cancer Inst.* **75**, 665 (1985).
3. A. G. Knudsen, *Cancer Res.* **45**, 1437 (1985).
4. B. Newman, M. A. Austin, M. Lee, M.-C. King, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 3044 (1988).
5. Four women with endometrial cancer (from families 1, 4, and 9) and two women with thyroid cancer (from families 5 and 13) were omitted entirely from the analysis. In family 3, two women have breast and ovarian cancer and are defined as affected at the ages of their breast cancer diagnoses; two other women died with ovarian cancer only and could not be included.
6. J. Ott, *Ann. Hum. Genet.* **47**, 311 (1983); G. M. Lathrop, J. M. Lalouel, C. Julier, J. Ott, *Am. J. Hum. Genet.* **37**, 482 (1985); N. Risch, *ibid.* **42**, 353 (1988).
7. D. E. Weeks and K. Lange, *Am. J. Hum. Genet.* **42**, 315 (1988).
8. L. Louie and M.-C. King, *ibid.* **41**, A174 (abstr.) (1987).
9. J. M. Hall *et al.*, *ibid.* **44**, 577 (1989).
10. A. P. Feinberg and B. Vogelstein, *Anal. Biochem.* **132**, 6 (1983); *ibid.* **137**, 266 (1984).
11. T. Maniatis, F. F. Fritsch, J. Sambrook, *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1982).
12. Y. Nakamura *et al.*, *Genomics* **2**, 302 (1988); J. Luty *et al.*, *Nucleic Acids Res.* **16**, 6250 (1988).
13. Centre d'Etude du Polymorphisme Humain, *CEPH Version 3 Database: Lod Scores and Recombination Estimates* (CEPH, Paris, 1989).
14. An example of the difficulties in typing VNTR alleles is presented by family 18, in which allele K is alleged to appear in cousins without pedigree evidence for identity by descent in intervening relatives. This assignment, and others like it, were made after analyzing several samples from the relatives on the same Southern blot, genotyping them without reference to their sample numbers or positions in the pedigree, and determining that the bands were indistinguishable among the four samples.
15. L. L. Cavalli-Sforza and M.-C. King, *Am. J. Hum. Genet.* **38**, 599 (1985).
16. For the 23 families as a whole, the method of affected pedigree members yields a T value of 1.13 ($P = 0.13$) for the T statistic unadjusted for allele frequencies, T equals 4.26 ($P < 0.001$) for the inverse square root weighting function, and T equals 5.49 ($P < 0.001$) for the reciprocal weighting function. The T statistic among the younger families was consistently significant, with T values of 2.60, 5.18, and 5.78, and empirical P values of 0.012, 0.001, and 0.003 for the three weighting functions, respectively.
17. P. Devilee *et al.*, *Genomics* **5**, 554 (1989); M. Genuardi, H. Tsihira, D. E. Anderson, G. F. Saunders, *Am. J. Hum. Genet.* **45**, 73 (1989); I. U. Ali, R. Lidereau, C. Theillet, R. Callahan, *Science* **238**, 185 (1987); C. Lundberg, L. Skoog, W. K. Cavenee, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 2372 (1987); C. Coles *et al.*, *Lancet* **336**, 761 (1990); T. Sato *et al.*, *Am. J. Hum. Genet.* **47**, A16 (abstr.) (1990).
18. M.-C. King, R. C. P. Go, R. C. Elston, H. T. Lynch, N. L. Petrakis, *Science* **208**, 406 (1980); M.-C. King *et al.*, *J. Natl. Cancer. Inst.* **71**, 463 (1983).
19. A. M. Bowcock, J. M. Hall, J. M. Hebert, M.-C. King, *Am. J. Hum. Genet.* **46**, 12 (1990); J. M. Hall *et al.*, *Genomics* **6**, 181 (1990).
20. B. Vogelstein *et al.*, *Science* **244**, 207 (1989); E. R. Fearon and B. Vogelstein, *Cell* **61**, 759 (1990).
21. V. A. McKusick, *Mendelian Inheritance in Man* (Johns Hopkins Univ. Press, Baltimore, ed. 10, 1990); E. Solomon and D. F. Barker, *Cytogenet. Cell Genet.* **51**, 319 (1989).
22. L. Coussens *et al.*, *Science* **230**, 1132 (1985); N. C. Popescu, C. R. King, M. H. Kraus, *Genomics* **4**, 362 (1989); P. P. Di Fiore *et al.*, *Science* **237**, 178 (1987); M. J. van de Vijver and R. Nusse, *Biochim. Biophys. Acta*, in press; P. Devilee and C. J. Cornelisse, *Cancer Surv.*, in press.
23. V. Luu-The *et al.*, *Mol. Endocrinol.* **4**, 268 (1990); G. Bevilacqua, M. E. Sobel, L. A. Liotta, P. S. Steeg, *Cancer Res.* **49**, 5185 (1989); A. M. Rosengard *et al.*, *Nature* **342**, 177 (1989); H. Roelink, E. Wagenaar, S. Lopes da Silva, R. Nusse, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 4519 (1990); M. Petkovich, N. J. Brand, A. Krust, P. Chambon, *Nature* **330**, 444 (1987); N. Brand *et al.*, *Nature* **332**, 850 (1988).

24. We thank Y. Nakamura, M. Litt, and D. Barker for probes; the CEPH for information on genetic markers; N. L. Petrakis and S. Wolman for help with interpretation of breast tumor pathology; J. Ott and N. Risch for statistical advice; E. Jones, S. Rowell, P. Zuppan, and J. Crandall for technical assistance; and E. Anderson and the Breast Cancer Task Force of NCI for encouragement. Supported by NIH grant CA27632.

12 September 1990; accepted 10 November 1990

AAAS–Newcomb Cleveland Prize

To Be Awarded for an Article or a Report Published in *Science*

The AAAS–Newcomb Cleveland Prize is awarded to the author of an outstanding paper published in *Science*. The value of the prize is $5000; the winner also receives a bronze medal. The current competition period began with the 1 June 1990 issue and ends with the issue of 31 May 1991.

Reports and Articles that include original research data, theories, or syntheses and are fundamental contributions to basic knowledge or technical achievements of far-reaching consequence are eligible for consideration for the prize. The paper must be a first-time publication of the author's own work. Reference to pertinent earlier work by the author may be included to give perspective.

Throughout the competition period, readers are invited to nominate papers appearing in the Reports or Articles sections. Nominations must be typed, and the following information provided: the title of the paper, issue in which it was published, author's name, and a brief statement of justification for nomination. Nominations should be submitted to the AAAS–Newcomb Cleveland Prize, AAAS, Room 924, 1333 H Street, NW, Washington, D.C. 20005, and **must be received on or before 30 June 1991**. Final selection will rest with a panel of distinguished scientists appointed by the editor of *Science*.

The award will be presented at the 1992 AAAS annual meeting. In cases of multiple authorship, the prize will be divided equally between or among the authors.